

## Machine Learning-Based Workflow for the Analysis of Project Portfolio

**Authors :** Jean Marie Tshimula, Atsushi Togashi

**Abstract :** We develop a data-science approach for providing an interactive visualization and predictive models to find insights into the projects' historical data in order for stakeholders understand some unseen opportunities in the African market that might escape them behind the online project portfolio of the African Development Bank. This machine learning-based web application identifies the market trend of the fastest growing economies across the continent as well skyrocketing sectors which have a significant impact on the future of business in Africa. Owing to this, the approach is tailored to predict where the investment needs are the most required. Moreover, we create a corpus that includes the descriptions of over more than 1,200 projects that approximately cover 14 sectors designed for some of 53 African countries. Then, we sift out this large amount of semi-structured data for extracting tiny details susceptible to contain some directions to follow. In the light of the foregoing, we have applied the combination of Latent Dirichlet Allocation and Random Forests at the level of the analysis module of our methodology to highlight the most relevant topics that investors may focus on for investing in Africa.

**Keywords :** machine learning, topic modeling, natural language processing, big data

**Conference Title :** ICMLDA 2018 : International Conference on Machine Learning and Data Analysis

**Conference Location :** Tokyo, Japan

**Conference Dates :** May 28-29, 2018